WO 01/90190 PCT/CA01/00763

SINGLE-DOMAIN ANTIGEN-BINDING ANTIBODY FRAGMENTS DERIVED FROM LLAMA ANTIBODIES

FIELD OF THE INVENTION

The invention relates to antigen-binding proteins, in particular to antigen-binding fragments of antibodies derived from a naïve library of llama antibodies and to a phage display library of such fragments. More particularly, the present invention relates to antigen-binding fragments of llama antibodies comprising at least a part of the variable heavy domain (VH or V_HH) of antibodies derived from a naïve library of llama antibodies and to a phage display library of such fragments.

BACKGROUND OF THE INVENTION

The immune system in vertebrates provides a defense mechanism against foreign intruders, such as foreign macromolecules or infecting microorganisms. The foreign invaders (antigens), both macromolecules (proteins, polysaccharides, or nucleic acids) and microbes (viruses or bacteria), are recognized through specific binding of the proteins of the host immune system to specific sites on the antigen surface, known as antigenic determinants.

As part of the immune system, B-cells of vertebrate organisms synthesize antigen-recognizing proteins known as antibodies or immunoglobulins (Ig). According to the clonal selection theory, an antigen activates those B-cells of the host organism that have on their surface immunoglobulins that can recognize and bind the antigen. The binding triggers production of a clone of identical B-cells that secrete soluble antigen-binding immunoglobulins into the bloodstream. Antibodies secreted by B-cells bind to foreign material (antigen) to serve as tags or identifiers for such material. Antibody-tagged antigens are then recognized and disposed of by macrophages and other effector cells of the immune system or are directly lysed by a set of nonspecific serum proteins collectively called complement. In this way a small amount of antigen

can elicit an amplified and specific immune response that helps to clear the host organism of the source of antigen. Through a complex process of gene splicing combined with additional mutation mechanisms, human B-cells have been estimated to produce a "library" (repertoire) of more than a billion (10^9) different antibodies that differ in the composition of their binding sites.

For most vertebrate organisms, including humans and murine species, their antibodies show a common structural pattern which consists of two identical light polypeptide chains and two identical heavy polypeptide chains linked together by disulfide bonds and numerous non-covalent interactions, resulting in a Y-shaped molecule. In humans, there are two different classes (isotypes), λ and κ, of the light chains, with no known functional distinction between them. The heavy chains have five different isotypes that divide immunoglobulins into five different functional classes (IgG, IgM, IgA, IgD, IgE), each with different effector properties in the elimination of antigen.

Of the above five classes, immunoglobulins of the IgG class are the major type in normal serum of humans and many other species and have the four-chain structure shown schematically in Fig. 1. Each chain of an IgG molecule is divided into domains of about 110 amino acid residues, with the light chains having two such domains and the heavy chains having four. Comparison of amino acid sequences between different IgGs shows that the amino-terminal domain of each chain (both light and heavy) is highly variable, whereas the remaining domains have substantially constant sequences. In other words, the light (L) chains of an IgG molecule are built up from one amino-terminal variable domain (VL) and one carboxy-terminal constant domain (CL), and the heavy (H) chains from one amino-terminal variable domain (VH) followed by three constant domains (CH1, CH2, and CH3).

The variable domains are not uniformly variable throughout their length. Three small regions of a variable domain, known as hypervariable regions (loops) or complementarity determining regions (CDR1, CDR2, and CDR3) show much more variability than the rest of the domain. These regions, which vary in size and sequence among various immunoglobulins, determine the specificity of the antigen-

antibody interaction. The specificity of an antibody of the type shown in Fig. 1 is determined by the sequence and size of six hypervariable loops (regions), three in the VL domain and three in the VH domain.

By partial digestion with papain, which cleaves the heavy chains in the hinge region, the IgG molecule can be broken down into two identical Fab fragments (Fragment, antigen binding) and one Fc fragment (Fragment, crystallizes easily). Each Fab fragment comprises one complete light chain (consisting of VL and CL domains) linked by a disulfide bridge and noncovalent interactions to a fragment of the heavy chain consisting of VH and CH1 domains. The Fc fragment comprises CH2 and CH3 domains from both heavy chains, also linked by disulfide bridges and noncovalent interactions. The part of the Fab fragment consisting of variable domains of the light and the heavy chain (VL and VH) is known as Fv fragment (Fragment, variable). In an Fv fragment, the variable domains VL and VH are not covalently bound. In an scFv (single chain Fv) fragment, the VL and VH domains are covalently linked by a short peptide linker (spacer), usually 15 to 20 amino acids long, introduced at the genetic level (see Fig. 2).

scFv fragments are recombinant fusion proteins and are produced by techniques of genetic engineering, by expressing in a suitable host, usually in bacteria, a chimeric gene coding for the fragment. Various other recombinant antibody fragments have been designed to substitute for large intact immunoglobulin molecules (see Fig. 2). Other than scFv fragments, these options include Fab or Fv fragments that are stabilized or covalently linked using various strategies (see, for example, Bird et al., *Science*, **242**, 423-426 (1988); Huston et al., *Proc. Natl. Acad. Sci. USA*, **85**, 5879-5883 (1988); Glockshuber et al., *Biochemistry*, **29**, 1362-1376 (1990); Jung et al., *Proteins*, 35-47 (1994); Reiter et al., *Biochemistry*, 5451-5459, 18327-18331 (1994); Young et al., *FEBS Lett.*, 135-139 (1995)). Small antigen-binding fragments of natural antibodies are advantageous for medical applications, for example cancer targeting and imaging, when small antigen-biding molecules are required to penetrate into solid tumors.

Recent advances in gene technology have greatly facilitated the genetic manipulation, production, identification and conjugation of recombinant antibody fragments and broadened the potential utility of antibodies as diagnostic and therapeutic agents. Of particular importance to such applications is the possibility to alter the fine specificity of the antibody binding site, to create small stable antigen-binding fragments, to prepare fusion proteins combining antigen-binding domains with proteins having desired therapeutic properties, for the purpose of immunotargeting, or to "humanize" antibodies of other species, for example murine antibodies (see Fig. 2).

The genetic engineering has also made possible to screen in vitro for antibodies having a predetermined binding specificity. This may be achieved by constructing first a gene library of antibodies or antibody fragments, for example by polymerase chain reaction (PCR)-amplification of cDNA derived from B-lymphocytes using suitable primers, or by in vitro gene synthesis. The gene library may contain sequences corresponding to certain fragments of natural antibodies, or randomized antigen-binding regions, or new combinations of heavy/light chains, thus creating the potential for generating antibodies which could never be obtained from natural sources, for example, antibodies to highly toxic substances or antigens tolerated by the human immune system. By random or designed mutations, the affinity or specificity of the antigen binding can be manipulated, for example, to reach affinities never observed with natural antibodies.

To screen a gene library, which may contain many millions or even billions of different clones, for genes of antibodies having the desired binding specificity, a selection system comparable to that of the immune system is required. Such a selection system can be achieved by inserting the library genes into the genome of microorganisms capable of displaying on their surface the antibody corresponding to the inserted gene, in analogy to the expression of an immunoglobulin antigen receptor on the surface of a B-cell. Microorganisms most frequently used for providing such a display are filamentous bacteriophages, such as fd or M13 phages (phage display). The collection of phage particles having inserted genes of a library of proteins, such as antibodies, and displaying these proteins on the particles'

surface is known as a phage display library. The display of the library of antibodies on the surface of phage particles provides a physical link between the antigen-binding function of an antibody and the antibody gene. Using the affinity to a preselected antigen, the whole organism (phage) displaying this affinity can be identified and separated out of billions of non-specific clones, usually through binding to the antigen immobilized on a support, technique usually referred to as panning (see, for example, Scott et al., *Science*, **249**, 386-390 (1990); Winter et al., *Annual Rev. Immunology*, **12**, 433-455 (1994)). Phage clones binding to the antigen can be then amplified and used to produce the specific antibody or antibody fragment in E. coli or in other suitable organism.

For naturally occurring antibodies, there are examples that whole heavy chains alone retain a significant binding ability in the absence of light chains. It is also well established, from structural studies, that the CDR3 of the heavy variable domain generally contributes the most to antigen binding, because CDR3 amino acid residues are responsible for most of the surface contact area and molecular interaction with the antigen (Padlan, E.A., *Mol. Immunology*, **31**, 169-217 (1984); Chothia et al., *J. Mol. Biol.*, **196**, 904-917 (1987); Chothia et al., *J. Mol. Biol.*, **186**, 651-663 (1985)). Less binding activity was observed for light chain. In view of these findings, attempts were made to isolate single VH domains. For example, VH domains were isolated from expression libraries derived from immunized mice (Ward et al., *Nature*, **341**, 544-546 (1989)). In another report, antigen-binding VH domains were rescued from an antibody phage library that was made from a vaccinated patient (Cai et al., *Proc Natl. Acad. Sci. USA*, **93**, 6280-6285 (1996)). Antigen-binding antibody fragments consisting of a single VH domain, known as dAbs or sdAbs (single-domain antibodies), are becoming an attractive alternative to single chain Fv (scFv) fragments. Despite smaller binding surface, their demonstrated affinity is comparable to that demonstrated by scFv fragments (Davies et al., *Biotech.*, **13**, 475-479 (1995)). Because of their smaller size, being half of the size of scFvs, sdAbs are amenable to detailed NMR structural studies (Davies et al., *FEBS Letters*, **339**, 285-290 (1994)). Additionally, due to their simpler structure, sdAbs are more stable and have simpler folding properties.

Recently, a new class of antibodies known as heavy chain antibodies (HCA, also referred to as two-chain or two-chain heavy chain antibodies) have been reported in camelids (Hamers-Casterman et al., *Nature*, **363**, 446-448 (1993); see also US 5,759,808; US 5,800,988; US 5,840,526; and US 5,874,541). Compared with conventional four-chain immunoglobulins of IgG-type, which are also produced by camelids, these antibodies lack the light chains and CH1 domains of conventional immunoglobulins. One of the salient features of these naturally occurring heavy chain antibodies is the predominant presence of Glu, Arg and Gly at VL interface positions 44, 45 and 47 (Kabat numbering), respectively, of their variable domain (designated V_HH). The same positions in the variable domain of the heavy chain of conventional four-chain antibodies (designated VH) are almost exclusively occupied by Gly, Leu and Trp. These differences are thought to be responsible for the high solubility and stability of camelid HCA variable domain (V_HH), as compared with the relative insolubility of VH domain of the conventional four-chain antibodies. Two more salient features of camelid V_HH domains are their comparatively longer CDR3 and high incidence of cysteine pairs in CDRs. It appears that cysteine pairs mediate the formation of a disulfide bridge and are therefore involved in modulating the surface topology of the antibody combining site. In the crystal structure of a camel sdAb-lysozyme complex, a rigid loop protruding from the sdAb and partly stabilized by a CDR disulfide linkage extends out of the combining site and penetrates deeply into the lysozyme active site (Desmyter et al., *Nature Struct. Biol.*, **3**, 803-811 (1996)).

More recently, a number of camelid sdAbs phage display libraries have been generated from the V_HH repertoire of camelids immunized with various antigens (Arbabi et al., *FEBS Letters*, **414**, 521-526 (1997); Lauwereys et al., *EMBO J.*, **17**, 3512-3520 (1998); Decanniere et al., *Structure*, **7**, 361-370 (1999)). By creating polyclonal libraries, many highly soluble sdAbs with high affinity and specificity have been isolated. However, it has been questioned whether sdAbs with desired affinity and defined conformations can be generated in the absence of prior immunization, i.e., with a naïve library (Lauwereys et al., supra). Immunization of domesticated valuable animals, such as camelids, raises serious ethical implications related to experiments with animals. Moreover, this approach has serious drawbacks because

most of the pathogenic antigens cannot be injected into camelids, as this could endanger their lives. Considering the above drawbacks and limitations of the prior art, there exists a strong need for the generation of phage display libraries of sdAb antibody fragments derived from naïve libraries of camelid antibodies, in particular sdAb fragments of camelid heavy chain antibodies, which libraries may become a universal source of sdAbs for *in vitro* selection against any antigen of interest as a target.

SUMMARY OF THE INVENTION

The present invention has overcome the above-discussed prior art limitations by generating a large size (in the order of 10^9) phage display library of antibody fragments of a non-immunized llama, which fragments comprise at least a part of the variable heavy domain (VH or V_HH domain) of llama antibodies. In a preferred embodiment, the fragments consist essentially of the variable heavy domain (VH or V_HH of llama antibodies (sdAb fragments). This library possesses a number of unique features which distinguish it from similar libraries generated from other camelids. The large size of the library considerably increases the probability of isolating therefrom antigen-binding fragments having high affinity to almost any predetermined target (antigen) of interest. This has been demonstrated by isolating from the library fragments binding specifically to several preselected antigens as targets.

Thus, according to one aspect, the invention provides a phage display library of antigen-binding fragments of llama antibodies, said fragments comprising at least a part of the variable heavy domain (VH or V_HH) of the antibodies. Preferably, the antigen-binding fragments consist of a complete variable heavy domain (VH or V_HH) of the antibodies (sdAb fragments)

According to another aspect, the invention provides an antigen-binding fragment of a llama antibody, said fragment comprising at least a part of the variable heavy domain

(VH or V_HH) of the antibody. Preferably, the antigen-binding fragment consists of a complete variable heavy domain (VH or V_HH) of the antibody (sdAb fragment).

According to yet another aspect, the invention provides a cDNA library comprising nucleotide sequences coding for antigen-binding fragments of llama antibodies, said library obtained by isolating lymphocytes from a biological sample obtained from a non-immunized llama; isolating total RNA from the lymphocytes; reverse-transcribing and amplifying RNA sequences coding for the antigen-binding fragments; cloning the amplified cDNA in a vector; and recovering the obtained clones. Preferably, the antigen-binding fragments consist of a complete variable domain (VH or V_HH) of the antibodies (sdAb fragment) and the cloning vector is a filamentous bacteriophage.

According to yet another aspect, the invention provides a process for the preparation of an antigen-binding fragment of a llama antibody, said fragment binding to a predetermined antigen, said process comprising the steps of isolating lymphocytes from a biological sample obtained from a non-immunized llama; isolating total RNA from the lymphocytes; reverse-transcribing and amplifying RNA sequences coding for antigen-binding fragments; cloning the cDNA sequences so obtained into a first vector, said first vector capable of a surface display of the corresponding antigen-binding fragments; subjecting the clones to antigen affinity selection and recovering clones having the desired affinity; for the recovered clones, amplifying DNA sequences coding for antigen-binding fragments; cloning the amplified DNA sequences into a second vector; transforming prokaryotic cells with the second vector under conditions allowing expression of DNA coding for antigen-binding fragments; and recovering the antibody fragments having the desired specificity.

Other advantages, objects and features of the present invention will be readily apparent to those skilled in the art from the following detailed description of preferred embodiments in conjunction with the accompanying drawings and claims.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a schematic representation of a typical four-chain IgG-type immunoglobulin (antibody) showing (a) the structure and arrangement of heavy and light chains and the approximate positioning of interchain disulfide bonds, and (b) the organization of the antibody molecule into paired domains.

Fig. 2 is a schematic representation of various modifications and fragments of IgG-type antibodies, and antigen-binding fusion proteins derived from such fragments.

Fig. 3 is a schematic representation of steps involved in construction of the phage display library of llama sdAb antibody fragments according to the present invention. For simplicity, only the coding sequences of the mRNA transcripts are shown. A, a: heavy chain mRNA of conventional four-chain (A) and two-chain heavy chain (a) antibodies; B, b: RT-PCR product derived from A and a, respectively; c: V_HH derived from heavy chain antibodies. Variable (VH) and constant (CH) domains are marked with dark and light shading, respectively.

Fig. 4 is a bar graph showing fractional occurrence of the CDR3 lengths. Gray bars represent data according to the present invention, whereas the white bars represent the published data for llama V_HH (Vu et al., *Mol. Immunol.*, **34**, 1121 – 1131 (1997)).

Fig. 5 is a graph showing global fitting to 1:1 interaction model of the binding of Yst9.1 scFv to immobilized Bruc.C6 sdAb fragment at 20, 100, 200, 300, 400, and 600 nM. Open circle lines represent experimental data points, whereas solid lines represent the fit.

Fig. 6 is a graph showing overlays of sensograms (A) and the Scatchard plot derived therefrom (B) for the binding of TNG.p1779 sdAb fragment (2.5 (f), 7.5 (e), 10 (d), 15 (c), 20 (b) and 30 (a) μM) to captured biotinylated p1779 peptide.

Fig. 7 is a graph showing the Scatchard plot derived from sensograms for the binding of TNG.PTH50 sdAb fragment to captured biotinylated PTH2 peptide.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

In the following, positions of amino acid residues in antibodies and antibody fragments are indicated according to the Kabat numbering.

The present invention provides a large size (in the order of 10^9) phage display library of single-domain fragments of variable heavy domains (VH and V_HH) of llama antibodies. The library, which has been generated using lymphocytes of a non-immunized animal (naïve library), can be used for *in vitro* selection against any antigen of interest as a target. The size of the library makes it highly probable that an antibody specific to the target will be identified among the library's sdAb fragments. This utility of the library has been demonstrated by isolating therefrom sdAbs binding specifically to various preselected antigens as targets.

The choice of a naïve library as the source of llama antibodies was based in part on the fact that the immune system of camelids has evolved over time in harsh environments and that its unique physiological and morphological features have helped the camelids to withstand water scarcity, adapt to climate extremes and develop a natural resistance to deadly viral diseases. The sero-epidemiological studies have confirmed that camelids produce antibodies to a great number of pathogenic viruses without developing the disease (Werney et al., Infectious Diseases of Camelids, Blackwell's Wissenschaft Verlag, Berlin (1995)). This means that antibodies of therapeutic importance can be isolated from the antibody repertoire of camelids without prior immunization with potentially dangerous pathogens or fragments thereof.

Another advantage of choosing a naïve library as the source of llama antibodies concerns anti-idiotypic antibodies. An anti-idiotypic antibody (a second antibody) recognizes the idiotope of another antibody (a first antibody) as an antigen, meaning that the first antibody recognizes in turn the second (anti-idiotypic) antibody as its antigen. Anti-idiotypic antibodies have gained a widespread clinical use, e.g., in vaccine development for cancer and cholera (Grant et al., *Clin. Cancer Res.*, 1319-

1323 (1999); Herlyn et al., *Ann. Med.*, 66-78 (1999); Maxwell-Armstrong et al., *Br. J. Surg.*, 149-154 (1998); Pierre et al., *Acta Gastroenterol. Belg.*, 430-436 (1992)) and in autoimmune disease therapy (Perosa et al., *Clin. Exp. Rheumatol.*, 201-210 (1997). They have also been shown to increase the protective immune response against parasites, bacteria and viruses (Feodorova et al., *J. Med. Microbiol.*, 751-756 (1999) and references therein). Since the original antigens (i.e., cancer, bacterial or viral antigens) may have been weakly- or non-immunogenic or toxic to the cells, anti-idiotypic antibodies have been used in their place to provide immune protection against diseases. However, in almost all cases reported to date, anti-idiotypic antibodies have been developed by immunization. The present invention eliminates the step of immunization and allows isolation of anti-idiotypic antibodies of potential diagnostic and therapeutic value from a naïve library.

Among the camelids, llama is the smallest animal which can survive in a severe, cold climate. Lymphocytes of a llama from a farm located in Osgoode (Canada) have been used to generate the phage display library of variable heavy domains of llama antibodies. From this library, sdAbs binding specifically to several preselected antigens have been subsequently isolated and characterized.

CONSTRUCTION OF A NAÏVE LLAMA sdAb PHAGE DISPLAY LIBRARY

Fig. 3 depicts a schematic representation of steps involved in the construction of the V_HH-derived sdAb phage display library. As the first step, lymphocytes from the fresh blood of llama (from a farm located at Osgoode, Ontario, Canada) were prepared and their RNA was isolated using techniques well known to those skilled in the art. RT-PCRs (reverse transcriptase-polymerase chain reactions) were performed using primers annealing at the 5' end of VH or V_HH and CH2 genes of IgG. The amplified products were separated and fragments of the expected size derived from conventional IgG (~900bp) and heavy chain IgG (~600bp) were observed on the agarose gel. The smaller fragment was gel purified and used in a second PCR to amplify the V_HH genes. The amplification products were cloned into fd-tet (GIIID) vector, between the leader signal and gene III, to produce fusion proteins, which were displayed on the filamentous phage particles using a modified procedure.

As is well known to those skilled in the art, the probability of isolating a protein with high affinity or specificity against a target (antibody) of interest increases with the size of the library. Generally, two different types of vectors are used for generating phage display libraries: phagemid vectors and phage vectors. Libraries having size in the order of 10^8 can be constructed with relative ease using phagemid vectors. However, a phagemid-based libraries suffers from some serious drawbacks. First, phagemid vectors provide typically a monovalent display and therefore may not select for lower binding (of lower affinity), but potentially important antibody fragments. Second, a phagemid-based library allows for the enrichment of phage particles displaying deleted versions of the antibody fragments. Such particles, often with no binding activity, are preferably selected during the panning process over those displaying the full-length fragments and therefore obscure the process of selection of the full-length binders. Third, constructing a phagemid-based library requires a helper phage and therefore library construction, panning and downstream phage binding assays become a far more complicated and tedious task. For these reasons the use a phage vector for the library construction is preferred.

One of the most widely used phage vectors is fd-tet (Zacher III et al., *Gene*, **9**, 127-140 (1980)) which consists of fd-phage genome, plus a segment of Tn10 inserted near the phage genome origin of replication. Tn10 contains a tetracycline resistance gene, tetA, and thus confers tetracycline resistance to the host cells carrying the fd-tet vector. It has often been observed that the size of the fd-tet based library was generally low (in the range of $10^5 - 10^6$) (Harrison et al., Methods in Enzymology [Ed. Abelson, J.N.], **267**, 83-109 (1996); Krebber et al., *FEBS Letters*, **377**, 277-331 (1995)), possibly due to the toxic effect of tetA gene product on the host cells. According to the modified procedure of the present invention, the library was propagated as plaques in the absence of tetracycline, resulting in a llama V_HH library of size of approximately $8.8x10^8$. This is the largest size library ever obtained using fd-tet vector. Due to its size, the library has an enhanced probability of selecting therefrom proteins (antibody fragments) binding to almost any given target (antigen).

It would be known to those skilled in the art that, at least in principle, the display library of the invention could be generated using vectors other than phages, such as bacteria (e.g., *E. coli*) (Daugherty et al., *Protein Eng.*, 613-621 (1999); Georgiou et al., *Nat. Biotechnol.*, 29-34 (1997)) or yeast (e.g., *Saccharomyces cerevisiae*) (Kieke et al., *Proc. Natl. Acad. Sci. USA.*, 5651-5656 (1999); Kieke et al., *Protein Eng.*, 1303-1310 (1997); Cho et al., *J. Immunol. Methods*, 179-188 (1998); Boder et al., *Nat. Biotechnol.*, 553-557 (1997)). Obtaining large libraries, comparable in size to phage display libraries, is, at least in theory, possible using these vectors. However, these display systems have not been of a widespread use, as they require expensive flow cytometry cell sorting instruments for selection. In addition, the *E. coli* display system is not suitable for panning against large macromolecules, such as proteins, due to the interference of the lipopolysaccharide layer of *E. coli* with the binding process (Boder et al., supra). Surface display of an scFv on mammalian cells has also been reported (Rode et al., *J. Immunol. Methods*, 151-160 (1999); Rode et al., *BioTechniques*, **650**, 652-656, 658 (1996)). However, no antibody library has been so far constructed using vectors other than phages, as the construction and screening in these alternative display systems are not as rapid or versatile as for phage display libraries.

SEQUENCE ANALYSIS

Colony PCR of 80 randomly selected clones showed that more than 60% had the full-length V_HH genes (sdAbs). The identity of the VL interface amino acids at position 44, 45 and 47 as well as the CDRs sequence of 28 randomly selected sdAbs have been determined and are summarized in Table 1. Fig. 4 shows the fractional occurrence of the CDR3 length. For comparison, previously published sequence data obtained from llama HCAs are also included. Similar to the previous results, the majority of the CDRs of the sequenced sdAbs are 13-17 amino acid long, demonstrating that the llama sdAb library of the invention is derived from heavy chain antibodies. However, the present library is distinct in several aspects from the known V_HH libraries.

Table 1. CDR/H1 sequences of 28 randomly selected dAbs from the llama library. The VL interface residues at positions 37, 44, 45 and 47 are also included. Position 35 is in each case the last residue in CDR/H1 sequence.

dAb	37	44	45	47	CDR1/H1		CDR2		CDR3	
C1	V	G	L	W	GFTFSSYYMS	SEQ ID NO: 1	GIYSDSSITAYADSVKG	SEQ ID NO: 29	MVMGPAATGYEY	SEQ ID NO: 57
C2	F	E	R	F	GRTFSNYHMG	SEQ ID NO: 2	SIKWSGGNTYYADSVKG	SEQ ID NO: 30	GSKYGGSWSRSQDAYNY	SEQ ID NO: 58
C4	F	E	R	F	GRIFSNAAMG	SEQ ID NO: 3	AIRWSDGNTYYADSVKG	SEQ ID NO: 31	GIGTFGSSWTRADRYRY	SEQ ID NO: 59
C5	Y	Q	R	L	RSIFSINTLG	SEQ ID NO: 4	WITSGGATYYADSMKG	SEQ ID NO: 32	RVPLDY	SEQ ID NO: 60
C7	F	E	R	F	GRSFSTYRVG	SEQ ID NO: 5	GINWNGVKTRYSDSMND	SEQ ID NO: 33	DQRFDGDDWSPSAFTR	SEQ ID NO: 61
C8	F	E	R	F	GNTISGYATG	SEQ ID NO: 6	AVTWSGYSVYYAKSPKG	SEQ ID NO: 34	VFVRTAGVPTLGEYDY	SEQ ID NO: 62
C9	F	G	R	F	GGSFSNYNMG	SEQ ID NO: 7	GIGWSGGRIIVADSVKG	SEQ ID NO: 35	TKQFFPLSN?SVWYDY	SEQ ID NO: 63
C12	W	K	R	F	GRIPRNYPIG	SEQ ID NO: 8	GISWTSGTTYFADSVKG	SEQ ID NO: 36	SERDFYTRNYYFTFESLYDY	SEQ ID NO: 64
C15	F	A	R	F	GESIASFNLG	SEQ ID NO: 9	AVSRTGETTDYADAVKG	SEQ ID NO: 37	DYNLGTEVTRKDSMYDF	SEQ ID NO: 65
C16	F	E	R	F	GRTFSSVSMG	SEQ ID NO: 10	AINWRGVSTYYADSVKG	SEQ ID NO: 38	RRNFEGNNSAGQYAY	SEQ ID NO: 66
C17	L	E	R	I	GLTFGDYAMG	SEQ ID NO: 11	TISRIGSTTYYADSVKG	SEQ ID NO: 39	SRYVLKYDKDAY	SEQ ID NO: 67
C22	F	E	R	F	GRTFSSVTMG	SEQ ID NO: 12	AMTRNSGSTYYADSVKG	SEQ ID NO: 40	KASMYGSTLYPPTGYNY	SEQ ID NO: 68
C24	F	E	R	F	GRTFSRFAMG	SEQ ID NO: 13	AISWSGGTTYGADSAKG	SEQ ID NO: 41	GRAVSDYDY	SEQ ID NO: 69
C25	Y	E	R	L	GSIFSESAMG	SEQ ID NO: 14	AITLDGRTNYAYYAEG	SEQ ID NO: 42	LRSRAVMDTIPNY	SEQ ID NO: 70
C26	F	E	R	F	GRTFSSDAMG	SEQ ID NO: 15	AISWSGGSTYYADSVKG	SEQ ID NO: 43	DRRRYYSGSYPPSEYDY	SEQ ID NO: 71
C29	V	G	L	W	GFTFSNFWMG	SEQ ID NO: 16	QINTGGDITTYSDSVKG	SEQ ID NO: 44	ARSVPLSDPRTYSS	SEQ ID NO: 72
C30	L	E	R	V	GRSFNHYIMG	SEQ ID NO: 17	SIDWNSGRTNYADSVKG	SEQ ID NO: 45	AAAASTLVGGSYDY	SEQ ID NO: 73
C31	Y	E	R	F	GLPFSTYSMG	SEQ ID NO: 18	VIGGGGNTYHAADSLKD	SEQ ID NO: 46	DRDFTIVAGFIRSQYSPRAVEY	SEQ ID NO: 74
C33	F	E	R	F	GRTFSTYTMG	SEQ ID NO: 19	AISRNSVGTYYRDSVKG	SEQ ID NO: 47	DPMYGRSVMSTRYNY	SEQ ID NO: 75
C34	F	D	R	F	GYTFSSHAMG	SEQ ID NO: 20	AISASGGNQYYKYFAKG	SEQ ID NO: 48	ATKQFSNAYSDYVHDYDY	SEQ ID NO: 76
C35	F	E	R	G	GFRFAEYAIG	SEQ ID NO: 21	YISTSDKTTYYSDFAEG	SEQ ID NO: 49	GLYYSDYRTPEYTEYVH	SEQ ID NO: 77
C40	F	E	R	F	GRTFSRFAMG	SEQ ID NO: 22	AISWSGGTAYGADSAKG	SEQ ID NO: 50	GRAVSDYDY	SEQ ID NO: 78
C43	V	G	L	W	GFTFVDYSMT	SEQ ID NO: 23	AINWNGRLTYYAESMKG	SEQ ID NO: 51	GELYGMGSKHDY	SEQ ID NO: 79
C44	V	G	L	W	GFTFSNYYMY	SEQ ID NO: 24	MVNTGGGGTRYADSVRG	SEQ ID NO: 52	DRPQSGWSMDY	SEQ ID NO: 80
C45	F	E	R	F	GLTFSSYVMG	SEQ ID NO: 25	AIITSGRSTYYADSVKG	SEQ ID NO: 53	TKWVVRRPADYNY	SEQ ID NO: 81
C46	F	E	R	F	GGTFTDYAMG	SEQ ID NO: 26	AINWGGYSTYYSDAVKG	SEQ ID NO: 54	DPQLITTPEYNY	SEQ ID NO: 82
C48	V	G	L	W	GFTFSNYYMY	SEQ ID NO: 27	MVNTGGGGTRYADSVRG	SEQ ID NO: 55	DRPQSGWSMDY	SEQ ID NO: 83
C49	F	E	R	F	GNTISDYATG	SEQ ID NO: 28	SIGRRTGWQVYSDSVKG	SEQ ID NO: 56	SQDSGFDTPVTESHLYGY	SEQ ID NO: 84

Previously generated camelid sdAb libraries were characterized by typical presence of Glu, Arg and Gly in positions 44, 45 and 47, respectively, of the VL interface of V_HH domain. The occurrence of cysteine at position 45 was also frequent in V_HH, as opposed to VH domain of four-chain IgGs. The present library, as shown by sequence analysis (Table 1), lacks these characteristics, as only one sdAb (C35) has Glu44, Arg45 and Gly47. The majority of sdAbs of the present library have Arg in position 45 of the VL interface. This occurrence of Arg45 is not unique to camelid V_HH, as a number of conventional antibodies, such as H1-I6 (VH) and V13 (VH), have been found to have Arg in position 45 (Blier et al., *J. Immunol.*, **139**, 3996-4006 (1987); Crews et al., *Cell*, **29**, 59-66 (1981)). The presence of Gly at position 35 was observed to always accompany Phe at position 37, unlike a previously reported llama library in which this pairing was observed in only 50% of the sequences. This is noteworthy in view of the fact that Gly at position 35 results in local conformational changes that allow Trp101 to stack with Arg45 in addition to engaging in aromatic–aromatic interactions involving Phe37 and Trp103. For the present library, 12 of 27 sdAbs have Trp at position 52a, whereas only 1 of the 51 previously published sequences have Trp at this position.

Another major difference between the present library and the previously reported V_HH libraries of camelids concerns the CDR cysteins. Previously generated libraries were characterized by a high incidence of cysteine pairs in CDRs, whereas none of the 28 sdAbs (Table 1) of the present library had any cysteine in their CDRs. The library of the invention is therefore characterized by a very low presence or by the absence of cysteine residues in CDRs.

Finally, the present library, which was designed and constructed to contain only antibody fragments consisting of variable heavy chain domains (V_HHs), also contains a substantial number of typical conventional variable heavy domains (VHs) (for example, sdAbs C1, C29, C43, C44 and C48 of Table 1, some sdAbs of Table 2). This contamination is most likely the results of PCR cross-overs between the VHs and V_HHs during the step of RT-PCR (Figure 3) (Tomlinson et al., *J. Mol. Biol.*, **227**, 776-798 (1992); Muyldermans et al., *Protein Eng.*, **7**, 1129-1135(1994)). These VHs are genuine antigen binding fragments, as shown in Table 2, produced in high yield

in *Escherichia coli*. They are highly soluble, have excellent temperature stability profiles and do not display any aggregation tendencies (Tanha et al., manuscript in preparation; Vranken et al., submitted). The very close similarity of these molecules to human VHs makes them potentially very useful as therapeutic sdAbs.

For the library of the invention, amino acids of the VL interface are most frequently:

at position 44 – Gly, Glu, Gln, Lys, Ala and Asp,

at position 45 – Leu, Phe, Pro and Arg, and

at position 47 – Trp, Tyr, Phe, Leu, Ile, Val and Gly.

For the library of the invention, CDRs can be selected from the following sequences:

CDR1/H1:	GFTFSSYAMS	(SEQ ID NO: 85)
	GFTFSSYYMS	(SEQ ID NO: 86)
	GFTFDEHAIG	(SEQ ID NO: 87)
	GFTVSSNHMT	(SEQ ID NO: 88)
	GFTFSSYHMA	(SEQ ID NO: 89)
	GFTFSRHQMS	(SEQ ID NO: 91)
	GFTFRTYYMN	(SEQ ID NO: 92)
	GFIFSSYAMS	(SEQ ID NO: 93)
	GFTFSTYAMT	(SEQ ID NO: 95)
	GFTFSGYAMS	(SEQ ID NO: 99)
	GFAFSNYRMT	(SEQ ID NO: 100)
	GFTFSRYAMS	(SEQ ID NO: 101)

CDR2:	GIEGGGGITRYADSVKG	(SEQ ID NO: 102)
	TIKPGGGSTYYADSVKG	(SEQ ID NO: 103)
	TIDIGGGRTYADSVKG	(SEQ ID NO: 104)
	RISSDGRNTYYADSVKG	(SEQ ID NO: 105)
	TINPGDGSTYYADSVKG	(SEQ ID NO: 106)
	HIDTGGSTWYAASVKG	(SEQ ID NO: 107)
	TINIDGSSTYYADSVRG	(SEQ ID NO: 109)
	GINSFGGSKYYADSVKG	(SEQ ID NO: 110)
	TINTSGRGTYYADSVKG	(SEQ ID NO: 112)
	AINSGGGSTSYADSVKG	(SEQ ID NO: 113)
	HIDTGGGSTWYAASVKG	(SEQ ID NO: 114)
	DINSGGDSTRNADSVKG	(SEQ ID NO: 115)
	SINSGGGSTYYADSVKG	(SEQ ID NO: 116)
	RINSIGDRISYADSVKG	(SEQ ID NO: 117)

CDR3:	AHGGYGAFGS	(SEQ ID NO: 119)
	YSGGALDA	(SEQ ID NO: 122)
	LSQGAMDY	(SEQ ID NO: 124)

IDRERAFTS	(SEQ ID NO: 127)
IDWERAFTS	(SEQ ID NO: 128)
QGYAGSYDY	(SEQ ID NO: 129)
LGVPGTFDY	(SEQ ID NO: 130)
TNRGIFDY	(SEQ ID NO: 131)
TPGSSGVYEY	(SEQ ID NO: 132)
TQTGSHDY	(SEQ ID NO: 133)
QVGTAYDY	(SEQ ID NO: 134)
RRGSSGVYEY	(SEQ ID NO: 135)

SELECTION AGAINST ANTIBODY ANTIGENS

Special cases of antibody-antigen reactions are those in which the antigen (Ag) is itself an antibody (Ab), as discussed above. Single domain anti-idioptypic (anti-Id) antibody fragments have been isolated from the library of the present invention using phage display technology and an antibody serving as antigen. Such anti-Id antibody fragments have great potential in both evoking the immune system responses to pathological antigens and in vaccine development.

Single chain Fv-Yst9.1 (anti-Brucella antibody)

The above-described naïve llama phage display library was panned against Yst9.1 scFv immobilized on micro-titer plates. A very high enrichment was observed in the case of anti-Brucella carbohydrate (Yst9.1 scFv), as all the 60 selected clones showed strong binding in phage ELISA to Yst9.1 scFv but no binding to the BSA control. Sequencing revealed 17 different sdAbs, some of which, were related to each other (Table 2). For example, Bruc.B3, Bruc.B10 and Bruc.C7.3 have the same CDR3. As another example, Bruc.C7.2, Bruc.D10 and Bruc.E6 have the same CDR3 in addition to the first two, which share the same CDR2. These common sequences were encoded by identical nucleotides raising the possibility that divergent sdAbs may have arisen as a result of PCR cross-over in vitro. Interestingly, the interface amino acids are generally Gly44, Leu45 and Trp47, typical of human/murine VH domain. In addition, none of the isolated sdAbs have any cysteine in CDR1, 2, or 3.

Table 2 also shows the identity of amino acids at positions 37, 44, 45 and 47 of the VL interface of V_HH domain. Interestingly, all sdAbs shown in the table have VL

Table 2. CDR/H1 sequences of dAbs which were isolated by panning the llama library against Yst9.1 scFv. The V_L interface residues at positions 37, 44, 45 and 47 are also included.

sdAb	37	44	45	47	CDR1/H1		CDR2		CDR3	
Bruc.B3	V	G	L	W	GFTFSSYAMS	SEQ ID No: 85	GIEGGGGITRYADSVKG	SEQ ID NO: 102	AHGGYGAFGS	SEQ ID NO: 119
Bruc.B10	V	G	L	W	GFTFSSYYMS	SEQ ID No: 86	TIKPGGGSTYYADSVKG	SEQ ID NO: 103	AHGGYGAFGS	SEQ ID NO: 120
Bruc.C7.3	F	G	F	S	GFTFDEHAIG	SEQ ID No: 87	TIDIGGGRTYADSVKG	SEQ ID NO: 104	AHGGYGAFGS	SEQ ID NO: 121
Bruc.B8	V	G	L	W	GFTVSSNHMT	SEQ ID No: 88	RISSDGRNTYYADSVKG	SEQ ID NO: 105	YSGGALDA	SEQ ID NO: 122
Bruc.D4.4	V	G	L	W	GFTFSSYHMA	SEQ ID No: 89	TINPGDGSTYYADSVKG	SEQ ID NO: 106	YSGGALDA	SEQ ID NO: 123
Bruc.C7.2	F	G	L	Y	GFTFDEHAIG	SEQ ID No: 90	HIDTGGSTWYAASVKG	SEQ ID NO: 107	LSQGAMDY	SEQ ID NO: 124
Bruc.D10	V	G	L	Y	GFTFSRHQMS	SEQ ID No: 91	HIDTGGSTWYAASVKG	SEQ ID NO: 108	LSQGAMDY	SEQ ID NO: 125
Bruc.E6	V	G	L	W	GFTFRTYYMN	SEQ ID No: 92	TINIDGSSTYYADSVRG	SEQ ID NO: 109	LSQGAMDY	SEQ ID NO: 126
Bruc.E3.1	V	G	L	W	GFIFSSYAMS	SEQ ID No: 93	GINSFGGSKYYADSVKG	SEQ ID NO: 110	IDRERAFTS	SEQ ID NO: 127
Bruc.E7.3	V	G	F	W	GFIFSSYAMS	SEQ ID No: 94	GINSFGGSKYYADSVKG	SEQ ID NO: 111	IDWERAFTS	SEQ ID NO: 128
Bruc.C6	V	G	L	W	GFTFSTYAMT	SEQ ID No: 95	TINTSGRGTYYADSVKG	SEQ ID NO: 112	QGYAGSYDY	SEQ ID NO: 129
Bruc.C5	V	G	P	W	GFTFSSYAMS	SEQ ID No: 96	AINSGGGSTSYADSVKG	SEQ ID NO: 113	LGVPGTFDY	SEQ ID NO: 130
Bruc.B7.1	V	G	L	Y	GFTFSRHQMS	SEQ ID No: 97	HIDTGGGSTWYAASVKG	SEQ ID NO: 114	TNRGIFDY	SEQ ID NO: 131
Bruc.B7.1A	V	G	P	W	GFTFSRYAMS	SEQ ID No: 98	DINSGGDSTRNADSVKG	SEQ ID NO: 115	TPGSSGVYEY	SEQ ID NO: 132
Bruc.D6	V	G	L	W	GFTFSGYAMS	SEQ ID No: 99	SINSGGGSTYYADSVKG	SEQ ID NO: 116	TQTGSHDY	SEQ ID NO: 133
Bruc.D5	L	G	F	W	GFAFSNYRMT	SEQ ID No:100	RINSIGDRISYADSVKG	SEQ ID NO: 117	QVGTAYDY	SEQ ID NO: 134
Bruc.F7	V	G	P	W	GFTFSRYAMS	SEQ ID No:101	DINSGGDSTRNADSVKG	SEQ ID NO: 118	RRGSSGVYEY	SEQ ID NO: 135

interface residues which are typical of murine or human VHs. For example, half of the sdAbs have Val37, Gly44, Leu45 and Trp47, which are highly conserved in murine and human VH. In addition, all sdAbs have Val37 and Gly44, and majority has Leu45 and Trp47. Six, three and one sdAbs are characterized by the presence of Phe45 or Pro45, Tyr45 and Ser45, respectively. It is interesting to note that the presence of the same VL interface residues in the conventional antibodies would render the isolated VH highly hydrophobic, resulting in their aggregation, which is not observed for llama antibodies.

With the presence of "human residues" at positions 37, 44, 45 and 47, the entire sequences of the Yst9.1-specific sdAbs are very homologous to human VH3 family sequences. A comparison of a consensus VH3 family sequence and the Yst9.1-specific sdAbs reveals amino acid differences at only five positions (Table 3). One of the five differences, the position 83 difference (Lys in the Yst9.1-specific sdAbs and Arg in the human consensus sequence) is conservative. Spatially, residues 6 and 108 are close and are located in the first and last (ninth) β-strands, respectively. The other three residues are positioned in non-CDR loops. Incorporation of some of these residues into an otherwise insoluble human V_H has rendered the domain soluble (unpublished results).

Table 3. Amino acid differences between a human VH3 family concensus sequence and the Yst9.1 binders listed in Table 2. Amino Acid positions are indicated in Kabat Numbers.

Amino acid position	6	74	83	84	108
V_HH	Ala	Ala	Lys	Pro	Gln
Human VH3 family	Glu	Ser	Arg	Ala	Leu

BINDING STUDIES

One of the anti-Yst 9.1 scFv sdAbs, Bruc.C6, was shown to be specific for its antigen by BIACORE analysis, as it bound to Yst 9.1 scFv (Fig. 5). The kinetic rate constants, k_a and k_d, obtained by the global fitting of the binding data, are shown in Table 5. The calculated K_d in this case is 380 nM (Table 5).

Table 5. Kinetic and equilibrium constants for the binding of Bruc.C6 to Yst9.1 ScFv and of TNG.P1779 to biotinylated peptide p1779. The values were determined from the retrospective sensograms and Scatchard plots in Figures 5, 6 and 7. ND = not determined.

	k_a ($M^{-1}s^{-1}$)	k_d (s^{-1})	K_d (M) (k_d/k_a)	K_d (M) (Scatchard plot)
Bruc.C6	1.4×10^4	5.5×10^{-3}	3.8×10^{-7}	ND
TNG.P1779	ND	ND	ND	1.1×10^{-5}

SELECTION AGAINST PEPTIDE ANTIGENS

These selection studies were carried out against peptides derived from granulin A and the parathyroid hormone (PTH).

A. GRANULIN A-DERIVED PEPTIDES

Proteins of granulin/epithelin family are thought to play a role in inflammation, wound repair, tissue modeling and regulating enzyme activity (Vranken et al., *J. Pept. Res.*, 590-597 (1999); Hrabal et al., *Nat. Struct. Biol.*, -752 (1996)). They are implicated as potential co-factors for HIV Tat protein and in modulating the growth of human epidermal carcinoma cells, and inhibition of their expression is known to inhibit the tumorigenecity of certain cells. The granulin motif has been found throughout the animal kingdom, in fish and insects, and encoded in the genome of a nematode worm. The motif consists of a parallel stacks of beta-hairpins pinned together by disulfide bonds. The structural sub-domain of granulin containing the first two beta-hairpin and spanning the first N-terminal 30 amino acids is also shared by growth factor proteins such as epidermal growth factors, transforming growth factor (TGF)-alpha, as well as the epithelial cell-specific TGF (TGF-e) which modulates the growth of human epidermal carcinoma cells. These growth factors interact with their receptors through their N-terminal beta-hairpin sub-domain and it is believed that epithelin/granulin family of proteins exert their growth modulating effect through the

same subdomain, by interacting with similar receptors. There have been continuous efforts in engineering stable sub-domains as possible drug candidates, with the aim of targeting specific proteins in vivo. The methodology has involved a rational amino acid substitution followed by assessing the effect of substitution on the stability of the sub-domain by NMR studies.

Solution structure of a 30-residue N-terminal sub-domain derived from carp granulin-1 has shown that the fragment forms two beta-hairpins similar to the one in the native protein. Unlike the carp granulin-1 sub-domain, the human counterpart (Tolkatchev et al., *Biochemistry*, 2878-2886 (2000); see also peptide p1779 in Table 6) was not stable outside the context of the native protein and a Q20P substitution (p1781) only slightly improved its stability. A substituted version incorporating D1V, K3H, S9I and Q20P, however, showed a well-folded stack of two beta-hairpins as in the carp granulin-1.

As an alternative and complement to NMR studies, antibodies can be used to probe the structural changes caused by amino acid substitution. The changes in the stability of a sub-domains brought about by amino acid substitutions may be manifested as changes in its affinity for an antibody probe compared to the wild type. Using peptides p1779, p1780 and p1781 shown in Table 6 as a model system it was demonstrated that a sdAb isolated from the llama sdAbs phage display library by panning against p1779 may serve as a structural probe. The sdAb binds to the p1779 peptide with a K_d of 10 μM, but shows no binding to the substituted versions of the peptide (peptides p1780 and p1781), which are known to have structures different from p1779. Other than serving as structural probes, such sdAbs can be used, for example, to interfere with granulin binding in pathways leading to cancer cell growth or HIV progression.

Table 6. Sequences of the human granulinA-derived peptide p1779 and its substituted versions p1780 and p1781. For panning experiments the peptides were labelled at the N-terminal through a $(Gly)_4$ linker.

Peptide	Sequence

Table 4. CDR/H1 sequences of dAbs which were isolated by panning the llama library against granuline A-derived peptides p1779 and p1781 (A) and PTH peptides (B). The V_L interface residues at positions 44, 45 and 47 are also included.

sdAb	44	45	47	CDR1/H1		CDR2		CDR3	
(A)									
TNG.P1779	Q	R	L	GSRRSFNVMG	SEQ ID No: 136	TITVGDTTSYAEAVKG	SEQ ID No: 158	EEWLGVRQNNY	SEQ ID No: 180
TNG.P1781-1	E	R	L	GDTFSINAYG	SEQ ID No: 137	AISGRGTNTFVADSVKG	SEQ ID No: 159	GEY	SEQ ID No: 181
TNG.P1781-2	G	L	W	GFTFRDYMMY	SEQ ID No: 138	SIYSDGSRTAYAASVKG	SEQ ID No: 160	MLLGPGARGYDY	SEQ ID No: 182
TNG.P1781-3	Q	R	L	GITFSEKHMA	SEQ ID No: 139	VITRGGTTNYGDSVKG	SEQ ID No: 161	DFYGLGFDY	SEQ ID No: 183
TNG.P1781-4	E	R	F	ERTFNSYAAA	SEQ ID No: 140	GITKNGVTYYAPSVTG	SEQ ID No: 162	APKYEGVSDTSSDYNY	SEQ ID No: 184
(B)									
TNG.PTH1	E	R	F	GRTFSSYGMG	SEQ ID No: 141	AMRESGADTHYADFVRG	SEQ ID No: 163	LDITTAASY	SEQ ID No: 185
TNG.PTH2	E	R	F	GRTFSSYGMG	SEQ ID No: 142	AMRESGADTHYADFVRG	SEQ ID No: 164	TINGAAR	SEQ ID No: 186
TNG.PTH4	K	R	L	GTSSGINAMV	SEQ ID No: 143	TITNSGKTDYAASAKG	SEQ ID No: 165	TINGAAR	SEQ ID No: 187
TNG.PTH5	E	R	F	GRTFSSYSMA	SEQ ID No: 144	AINWRSSVTAYADSVKG	SEQ ID No: 166	EALPGTYGLDY	SEQ ID No: 188
TNG.PTH7	Q	R	L	VSTFSIGAIG	SEQ ID No: 145	GISGGGSTYYTDSVKG	SEQ ID No: 167	ILAGGLLAF	SEQ ID No: 189
TNG.PTH8	Q	R	L	GSTFSGNDIG	SEQ ID No: 146	VISDGGYTSYATSVKG	SEQ ID No: 168	GGSSGTF	SEQ ID No: 190
TNG.PTH9	E	R	F	GRTFSSYGMG	SEQ ID No: 147	AISWGAGTPYYADSVKG	SEQ ID No: 169	TINGAAR	SEQ ID No: 191
TNG.PTH10	E	R	I	GRTFSDIAMA	SEQ ID No: 148	AIDWNGGTTYYTTFVKG	SEQ ID No: 170	LDITTAASY	SEQ ID No: 192
TNG.PTH11	E	R	F	GQTLNTYVMG	SEQ ID No: 149	AINWRDTSTYYQDSVKG	SEQ ID No: 171	TINGAAR	SEQ ID No: 193
TNG.PTH12	E	R	F	GPTSITYGMA	SEQ ID No: 150	AVTPSGGAAAYADSVKG	SEQ ID No: 172	GTELAPKTATGA	SEQ ID No: 194
TNG.PTH14	E	R	F	GGDVSTYAMV	SEQ ID No: 151	LLSRSGRTTNYADSVKG	SEQ ID No: 173	GSN	SEQ ID No: 195
TNG.PTH15	Q	R	L	GRTFGSYTMG	SEQ ID No: 152	RINSAGRTMYADSVKG	SEQ ID No: 174	GTVLSVATGPYGY	SEQ ID No: 196
TNG.PTH18	E	R	F	GRTFSSYGMG	SEQ ID No: 153	SINWRGSSTYYADSVKG	SEQ ID No: 175	WGAGEDEDY	SEQ ID No: 197
TNG.PTH22	Q	R	L	GSLSRITVMG	SEQ ID No: 154	IITSSGGTDYADSVKG	SEQ ID No: 176	KSRDSAGLSWDY	SEQ ID No: 108
TNG.PTH23	Q	R	V	GSISSFDAMA	SEQ ID No: 155	IITSGGATNYADSVKG	SEQ ID No: 177	LVASTVTSSVS	SEQ ID No: 199
TNG.PTH50	E	R	F	GRPFSSFAMG	SEQ ID No: 156	AISASGGETYYTGSLKG	SEQ ID No: 178	TINGAAR	SEQ ID No: 200
TNG.PTH61	E	R	F	GRTFSSYHMG	SEQ ID No: 157	AINWSGDTTYYEASVKG	SEQ ID No: 179	QTRPRPYGTSRAEGDYGY	SEQ ID No: 201

p1779	DVKCDMEVSCPDGYTCSRLQSGAWGCSPFT	SEQ ID No: 202
p1780	VVHCDMEVICPDGYTCSRLPSGAWGCSPFT	SEQ ID No: 203
P1781	DVKCDMEVSCPDGYTCSRLPSGAWGCSPFT	SEQ ID No: 204

Human granulin A-derived peptides

Solution panning was performed against human granulin A-derived peptide, p1779, and its substituted versions, p1780 and p1781 (Table 6). After four rounds of panning against p1779, phage sdAbs from all 48 clones tested were shown to bind to the target antigen. In the case of p1781, only eight binders (four different sequences, Table 4) were identified. No binder was identified for p1780 even after fifth round and performing the panning experiment two more times under different conditions.

Sequencing of twenty-one p1779-specific sdAb genes identified one fragment, namely, TNG.P1779, which was further expressed for detailed binding studies by BIACORE. In agreement with the phage ELISA results, TNG.P1779 was shown to be active by BIACORE analysis in which biotinylated p1779 was captured on a SA-coated CM5 sensor chip (Fig. 6, part A). No binding was detected to the reference surfaces on which a similar amount of p1780 or p1781 had been captured (data not shown). A Scatchard plot of the binding data gave a K_d of 1.1×10^{-5} M (Table 5). These results demonstrate that the TNG.P1779 behaves like a structural probe, sensing the structural changes, which occur in p1780 or p1781 as a result of amino acid substitutions.

B. PARATHYROID HORMONE-DERIVED PEPTIDE

Parathyroid hormone (PTH) is the major regulator of serum calcium levels and its use for the treatment of bone loss due to osteoporosis has been postulated. Osteoporosis, which is characterized by bone loss, strikes at any age, affects both men and women, although women with higher frequency, and can results in

hospitalization, disability and death (Morley et al., *Current Medicinal Chemistry*, **6**, 1095-1106 (1999); Whitfield et al., *Drugs & Aging*, **15**(2), 117-129 (1999)). Most of the available drugs slow down or stop further bone loss, but have no bone growth-stimulating effects, hence are not capable of replacing the lost bones.

The bone-building action of the parathyroid hormone (PTH) and its implications for the treatment of osteoporosis has been recently reviewed (Whitfield et al., supra). PTH is expressed as a 115 amino acid precursor and secreted as a 84-residue peptide, but its bone growth-stimulating effects have been related to its N-terminal 34-residues peptide and shown to be the case in human trials. More recently, mutated and cyclized PTH peptide analogues have been shown to be more potent bone growth stimulators in in vitro studies (Morley et al., *Expert Opin. Therap. Pat.*, **8**, 30-37 (1998)). These analogues, which have been patented, show great promise as drugs for the treatment of osteoporosis and are currently at the clinical trial stage. However, to meet the regulatory requirements, the pharmacokinetics of these drugs needs to be monitored following their administration to human subjects. This can be achieved by obtaining reagents, such as antibodies, capable of specifically recognizing the PTH analogues present in biological samples. Within the past two years, attempts were made to raise such antibodies by conventional hybridoma technology, but no success was reported. In the present study, a number of sdAbs specific for the PTH peptide analogues PTH1 and PTH2 (Table 7) have been isolated from the phage display library of sdAb fragments of heavy chain antibodies derived from a naïve library of llama antibodies.

Table 7. Sequences of PTH1 and PTH2 peptides corresponding to N-terminal residues 17-31 (PTH1) and 1-31 (PTH2) of the human PTH. Compared to the human PTH, these analogs have a substitution at positions 37(K37L) and a β-lactam bond connecting the side chains of ^{22}E and ^{26}K.

Peptide	Sequence	
PTH1	^{17}SMERVEWLRKLLQDV^{31}	SEQ ID No: 205
PTH2	^{1}SVSEIQLMHNLGKHLNSMERVEWLRKLLQDV^{31}	SEQ ID No: 206

Human PTH-derived peptides

Panning against PTH1 resulted in the identification of thirteen different sdAbs, all of which bound to PTH1 in a phage ELISA (Table 4, TNG.PTH1 through TNG.PTH18). Four binders were identified for PTH2 (Table 4, TNG.PTH22, TNG.PTH23, TNG.PTH50 and TNG.PTH61). The binding sdAbs were expressed and purified in large quantities. The expression level was high and for one particular sdAb it exceeded 200 mg of protein per liter of bacterial culture. Three sdAbs were characterized in more details by surface plasmon resonance and shown to bind to their target antigens (Table 8). Fig. 7 shows the binding profile for TNG.PTH50 which was isolated by panning against PTH2. The calculated K_d for TNG.PTH50 is 4.3×10^{-6} which is shown in Table 8.

Table 8. Equilibrium constants for the binding of TNG.PTH22, TNG.PTH23 and TNG.PTH50 to biotinylated PTH2. The values were determined from the respective sensograms and Scatchard plots, as shown in Fig. 7 for TNG.PTH50.

sdAb	K_d (M)
TNG.PTH22	1.4×10^{-5}
TNG.PTH23	5.7×10^{-5}
TNG.PTH50	4.3×10^{-6}

EXPERIMENTAL

All reagents were chemical grade purchased from various companies. Unless stated otherwise, the media were prepared as described (Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbour Laboratory, Cold Spring Harbour, NY (1989)). Phosphate-buffered saline (PBS) was prepared as described

(Sambrook et al., supra). Induction medium was the same as Terrific Broth except that it contained no salts. Agarose top was prepared by combining the following reagents in a total volume of 1 liter: 10 g Bacto-tryptone, 5 g yeast extract, 10 g NaCl, 1 g MgCl2.6H2O, and 7 g agarose. The mixture was autoclaved and stored solid at room temperature. The oligonucleotides were synthesized using the Applied Biosystems 394 DNA/RNA synthesizer. DNA sequencing was performed by the dideoxy method (Sanger et al., Biotechnology, 104-108 (1992)) using the AmpliTaq DNA Polymerase FS kit and 373A DNA Sequencer Stretch (PE Applied Biosystems, Mississauga, ON, Canada). The host bacteria used for cloning was TG1: *supE hsd*5 *thi* Δ(*lac-proAB*) F' [*traD*36 *proAB*$^{+}$ *lacI*^{q} *lacZ*ΔM15]. All the cloning steps were performed as described (Sambrook et al., supra). The vector fd-tet was purchased from American Type Culture Collection (Manassas, VA) and engineered such that it contained ApaI and NotI restriction sites immediately following the gIIIp leader sequence codons (Simon J. Foote, personal communications).

Construction of naive llama sdAb library

Total RNA was isolated from the leukocytes of freshly-drawn heparinized blood of a male Llama (Lama glama) using QIAamp RNA Blood MiniTM kit (QIAGEN, Mississauga, ON, Canada) and following the recommended protocol. The concentration of RNA was calculated by measuring the A260 value and assuming 1 A260 = 40 µg/ml. Reverse transcription-polymerase chain reaction (RT-PCR) was performed on a total of 5.3 µg RNA using the HotStarTaq PolymeraseTM kit (QIAGEN). The primers used included a CH2-specific primer, LlamaFOR, 5'(CGCCATCAAGGTACCAGTTGA)3' [SEQ ID No: 207] and LlamaBACK primer, 5'(GATGTGCAGCTGCAGGCGTCTGGRGGAGG)3' [SEQ ID No: 208], which anneals to the 5' flanking region of VH genes. Amplified product of approximately 600 base pair was purified from the agarose gel using QIAquick Gel ExtractionTM kit (QIAGEN) and subjected to a second round of PCR using the primers LlamaApaII, 5'(CATGACCACAGTGCACAGGAKGTSCAGCT)3' [SEQ ID No: 209] and LlamaNotI, 5'(CGATTCTGCGGCCGCTGAGGAGACGGTGACCTG)3' [SEQ ID No: 210]. The PCR mixture contained 10 pmol/µl each of the two primers, 1X buffer (Perkin Elmer), 200 µM each of the four dNTPs and 0.05 unit/µl AmpliTaqTM DNA polymerase (Perkin Elmer). PCR protocol consisted of an initial denaturation step at 95°C for 15

min followed by 35 cycles of 94°C for 30 sec, 45°C for 30 sec, and 72°C for 1 min, and a final extension step at 72°C for 10 min. The primers were complimentary to the 5' and 3' ends of the amplified product and incorporated ApaII and NotI restriction sites (underlined) at the end of VH genes. The amplified products were purified using QIAquick PCR Purification kitTM (QIAGEN), cut sequentially with ApaII and NotI restriction endonucleases, purified again, ligated to the ApaII/NotI-treated fd-tet phage vector and desalted using the above kit. Electrocompetent TG1 cells were prepared (Tung et al., *Trends Genet.*, 128-129 (1995)) and 1.5 µg of the ligated product was mixed with 40 µl of competent E. coli strain TG1 and the cells were transformed by electroporation using the BIO-RAD Gene PulserTM according to the manufacturer's instructions. The transformed cells were immediately transferred into 1 ml of SOC medium and split into 3 sterile tubes containing 3 ml of 50°C agarose top, vortexed immediately, poured onto pre-warmed 2xYT petri dishes, and incubated at 37°C overnight. The phage particles were eluted by adding five ml of sterile PBS to the plates gently shaked at 4°C for 3 hr. The phage-containing PBS was collected, the plates were rinsed with an additional 5 ml PBS and the two supernatants were combined in a centrifuge bottle. The contents were centrifuged at 6000g for 15 min at 4°C, the supernatant was decanted into a sterile centrifuge bottle and the phage was purified as described (Harrison et al., supra). At the end of the purification, the phage pellet was dissolved in 20 ml of sterile PBS and stored in liquid nitrogen in 100 µl aliquots.

To determine the size of the library, immediately following the transformation and after the addition of the SOC medium, a small aliquot of the electroporated cells was serially diluted in exponentially growing TG1 cells. 200 µl of the diluted cells was mixed with 3 ml of 50°C agarose top and immediately poured onto 2xYT plates pre-warmed to 37°C. Plates were incubated overnight at 37°C and the number of plaques was used to determine the size of the library.

Panning

Panning was performed using the Nunc-Immuno MaxiSorpTM 8-well strips (Nunc). Briefly, the wells were coated overnight by adding 150 µl of 100 µg/ml antigen in

PBS. In the morning, the wells were rinsed three times with PBS and subsequently blocked with 400 µl PBS-2% (w/v) skim milk (2% MPBS) at 37°C for 2 hr. The wells were rinsed as above and 1012 transducing units phage in 2% MPBS were added. The mixture was incubated at room temperature for 1.5 hr after which the unbound phage in the supernatant was removed. The wells were rinsed 10 times with PBS-0.1% (v/v) Tween 20 and then 10 times with PBS to remove the detergent. The bound phage was eluted by adding freshly prepared 200 µl 100 mM triethylamine, pipetting the content of the well up and down several times and incubating the mixture at room temperature for 10 min. The eluted phage was transferred to a tube containing 100 µl 1 M Tris-HCl, pH 7.4 and vortexed to neutralize the triethylamine. Following this, 10 ml of exponentially growing TG1 culture was infected with 150 µl eluted phage by incubating the mixture at 37°C for 30 min. Serial dilutions of the infected cells were used to determine the titer of the eluted phage as described in the previous section. The remainder of the infected cells was spun down and then resuspended in 900 µl 2xYT. The cells were mixed in 300 µl aliquots with 3 ml agarose top and the phage propagated on the plates overnight at 37°C. In the morning the phage was purified, the titer was determined, and a total of 10^{11} transducing units phage were used for further rounds of selection.

Solution Panning

Solution panning was performed using SA-PMP (1 mg/ml) obtained from Promega (Madison, WI). To maintain SA-PMP in solution during the panning process, the reaction tubes were flicked frequently during the incubation period. Briefly, for each target antigen 2 x 100 µl SA-PMPs was first dispersed by gently flicking the bottom of the tubes, and then captured at the side of the tube in a magnetic stand (approximately 30 sec.) followed by careful removal of the supernatant. SA-PMPs were re-suspended in 100 µl 1X PBS, re-captured and the supernatant was removed. This washing process was repeated three times. To remove any possible streptavidin binders from the phage library the phage particles were pre-incubated with SA-PMP in 2% MPBS for 1 hr at room temperature and the magnetic beads were captured. To form the phage-antigen complex, 10^{12} t.u. phage (10^{11} t.u. for further rounds) in the supernatant was incubated in 2% MPBS containing 20 mg/ml

BSA, 0.05% Tn20 and 1 μg/ml biotinylated antigen in a total volume of 150 μl for 1 hr at room temperature. In a second tube 100 μl of the washed SA-PMP was blocked in 400 μl 2% MPBS at 37oC for 2 hr. The supernatant was discarded and the phage-biotinylated antigen complex solution from the first tube was added to the blocked SA-PMP at room temperature for 30 min. The supernatant was removed and the complex-bound SA-PMPs were washed twice with 100 μl PBS and then once with 100 μl 2% MPBS containing 0.05% Tn 20; this sequence of washes was repeated another three times and then finally SA-PMPs were washed twice with PBS. The bound phage was eluted by adding 200 μl of 100 mM freshly prepared triethylamine and standing at room temperature for 10 min. Phage elution, propagation, titering and purification were performed as described for solid phase panning. Depending on the antigen for the final third and fourth rounds the procedure preceding the elution step was modified as described below. Following the initial washing step, 100 μl SA-PMPs were blocked followed by removal of supernatant and subsequent incubation of SA-PMPs with 100 μl of 5 μg/ml biotinylated antigens in 2% MPBS at room temperature for 30 min. The antigen-bound SA-PMPs were washed 5 times with 0.5 % MPBS and then incubated with phage in 2% MPBS at room temperature for 1.5 hr in a total volume of 100 μl. The supernatant was removed and the phage bound SA-PMPs were washed eight times with 0.5% MPBS and two times with PBS before proceeding with the elution step.

Phage enzyme-linked immunosorbent assay (phage ELISA)

Individual phage-infected TG1 colonies were used to inoculate 200 μl of LB in sterile 96-well plates. The cells were grown overnight at 100 rpm and 37°C. In the morning, the plates were spun down in a bench top centrifuge, and the sdAb phage-containing supernatant was used for phage ELISA as described below. Briefly, Nunc-Immuno MaxiSorpTM plates (Nunc) were coated overnight at 4°C with 150 μl of 10 μg/ml of target antigen or control proteins in PBS. The contents were removed and the plates were tapped on a paper towel to remove any liquid remaining in the wells. The wells were blocked by adding 300 μl of PBS-2% (w/v) skim milk (2% MPBS) and incubating for 2 hr at 37°C. The contents of the wells were emptied as before, 100 μl of sdAb phage supernatant in 2% MPBS was added, and the wells

were incubated at room temperature for 1.5 hr. For biotinylated antigen, the plates were pre-coated with 5 µg/ml streptavidin overnight followed by blocking. The wells were then coated with the target antigen by incubating plates with 150 µl of I µg/ml biotinylated antigen at room temperature for 30 min. The wells were washed 5x with PBS-0.05% (v/v) Tween 20 (PBST) and then incubated with phage. For control experiments no coating with the biotinylated antigen was performed. The contents were emptied again and the wells were washed 5 times with PBST and subsequently blotted on a paper towel to remove any remaining wash buffer. 100 µl of the recommended dilution of HRP/Anti-M13 Monoclonal Conjugate (Amersham Pharmacia Biotech, Montreal, QC, Canada) in 2% MPBS was added and the wells were incubated at room temperature for 1 hr. The wells were washed six times as before and the binding of sdAb to the antigen was detected colorimetrically by adding 100 µl of equal mixtures of TMB Peroxidase Substrate and H2O2 (KPL, Maryland, USA) at room temperature for several min. The reaction was stopped by adding 100 µl of 1 M H3PO4 and the A450 was measured by DYNATECH MR5000 ELISA reader (DYNATECH).

Sub-cloning and expression of sdAbs

sdAb genes were amplified out of the phage vector by PCR using the primers, VH.Bbs, 5'(TATGAAGACACCAGGCCGATGTGCAGCTGCAGGCG)3' [SEQ ID No: 211], and VH.Bam, 5'(TATGGATCCTGAGGAGACGGTGACCTG)3' [SEQ ID No: 212] which also introduced BbsI and BamHI sites at the ends of the amplified fragments. sdAb genes were subsequently purified, cut sequentially with BbsI and BamHI restriction endonucleases, purified again with QIAquick Gel ExtractionTM kit (QIAGEN), and ligated to the BbsI/BamHI-treated pSJF-2 vector. An aliquot of the ligated product was used to transform E. coli strain TG1. Transformants were selected on ampicillin plates and the clones harbouring the sdAb genes were identified by PCR and sequencing. For expression, single positive clones were used to inoculate 25 ml of LB containing 100 µg/ml ampicillin and the culture was shaken at 240 rpm at 37°C overnight. In the morning, the entire overnight culture was used to inoculate 1 liter of M9 medium supplemented with 5 µg/ml vitamin B1, 0.4% casamino acid and 100 µg/ml ampicillin. The culture was shaken at room

temperature for 30 hr at 180 rpm and subsequently supplemented with 100 ml of 10X induction medium and 100 µl of 1 M isopropylthio-D-galactoside. The culture was shaken for another 60 hr, the periplasmic fraction was extracted by osmotic shock (Anand et al., *Gene*, 39-44 (1991) and the presence of sdAb in the extract was detected by Western blotting (MacKenzie et al., *Biotechnology N.Y.*, 390-395 (1994)). The periplasmic fraction was dialyzed extensively in 10 mM HEPES (N-[2-hydroxyethyl]piperazine-N'-[2-ethanesulfonic acid]) buffer pH 7.0, 500 mM NaCl. The presence of the sdAb C-terminal His5 tag allowed a one step protein purification by immobilized metal affinity chromatography using HiTrap ChelatingTM column (Phamacia). The 5-ml column was charged with Ni^{2+} by applying 30 ml of a 5 mg/ml NiCl2.6H2O solution and subsequently washed with 15 ml deionized water. Purification was carried out as described (MacKenzie, supra) except that the starting buffer was 10 mM HEPES buffer, 10 mM imidazole, 500 mM NaCl, pH 7.0, and the bound protein was eluted with a 10-500 mM imidazole gradient. The purity of the protein was determined by SDS-PAGE (Laemmeli U.K., in: Proteases and biological control [Reich et al., ed.], Cold Spring Harbour Laboratory, pp. 661-676 (1975)). sdAb preparation was further subjected to gel filtration chromatography using Superdex 75 column (Pharmacia) as described (Deng et al., *Proc. Natl. Acad. Sci. USA.*, 4992-4996 (1995)) and the purified monomer species were used in binding studies by surface plasmon resonance.

Surface Plasmon Resonance Analysis

Binding studies were performed using BIACORE 1000 (Jonsson et al., *BioTechniques*, 620-627 (1991)) available from Biacore Inc., Piscataway, NJ. Binding of the anti-Yst9.1 sdAbs to Yst9.1 scFv was assessed under the same conditions except that in this case sdAb was immobilized (540 RU) and the flow rate was set at 20 µl/min. For PTH binders 186 RU (PTH2) or 70 RU (control peptide) was immobilized and the flow rate was also set at 20 µl/min. Surface regeneration was achieved by washing the sensor chips with HBST buffer. In the case of p1779 binder, sdAb was passed over biotinylated p1779 (520 RU) or p1780 and p1781 control peptides (420 RU) which had been captured on a CM5 sensor chip coated with streptavidin (2260 RU). Kinetic rate constants were determined using

BIAevaluation software and fitting to 1:1 interaction model. Affinity constants were calculated from the kinetic rate constants and by Scatchard analysis of equilibrium binding data as described (MacKenzie et al., *J. Biol. Chem*, 1527-1533 (1996)).

Although various particular embodiments of the present invention have been described hereinbefore for the purpose of illustration, it would be apparent to those skilled in the art that numerous variations may be made thereto without departing from the spirit and scope of the invention, as defined in the appended claims.